

1. Visualization **One**

For this visualization, a scatter plot was picked to compare between two quantitative variables, airline delay and distance between origin and destination airports. A scatter plot was used to highlight the relationship between the two variables, which was then found to be linear; conveying that as the distance travelled increases, the amount of delay also increases.

A marking of size was used to represent the amount of trip cancellations made by each airline. Size was picked because it was easier to see patterns using it without overcomplicating the visualization.

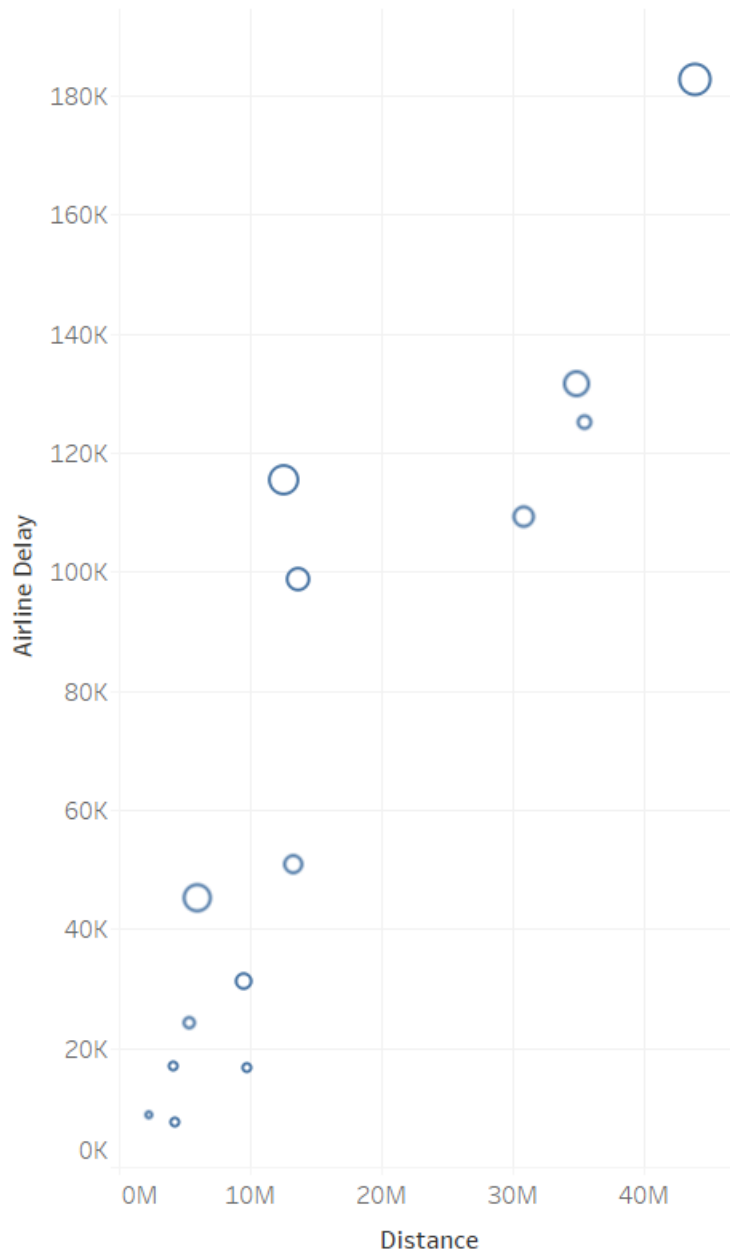
It was found that most cancellations occurred in airlines with average to high distance travelled and airline delay.

A filter with airline names was used to be able to answer questions like:

What was the sum of the airline delay encountered with respect to the sum of distances between origin and destination airports for each flight for a specific airline?

Before Filtering:

Airline Delays vs. Distance Travelled



Airline

- ☒ (All)
- ☒ AA
- ☒ AS
- ☒ B6
- ☒ DL
- ☒ EV
- ☒ F9
- ☒ HA
- ☒ MQ
- ☒ NK
- ☒ OO
- ☒ UA
- ☒ US

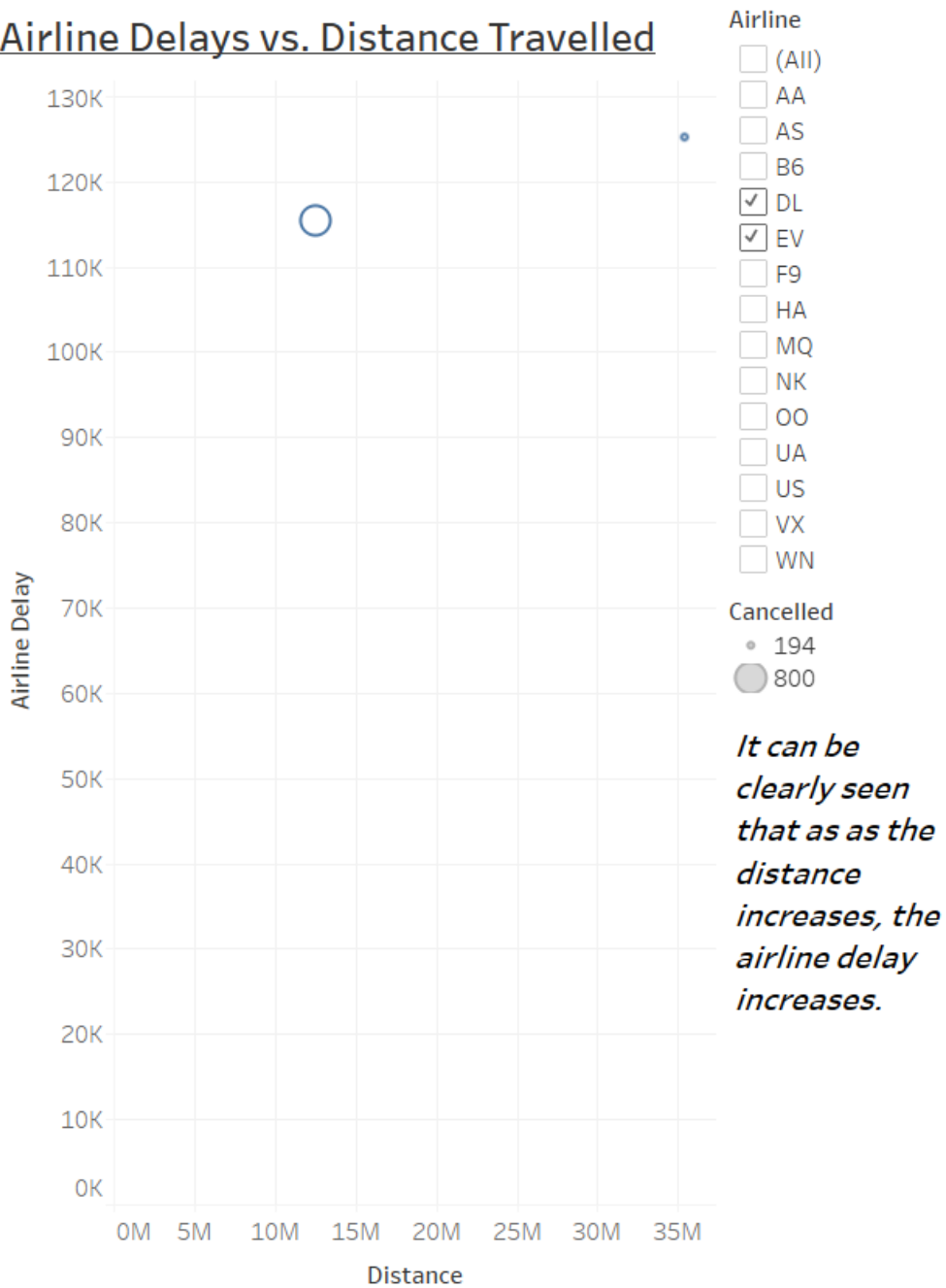
Cancelled

- 8
- 22
- 102
- 194
- 202
- 233
- 351
- 449
- 487
- 723
- 800
- 818

It can be clearly seen that as the distance increases, the airline delay increases.

After Filtering:

Airline Delays vs. Distance Travelled



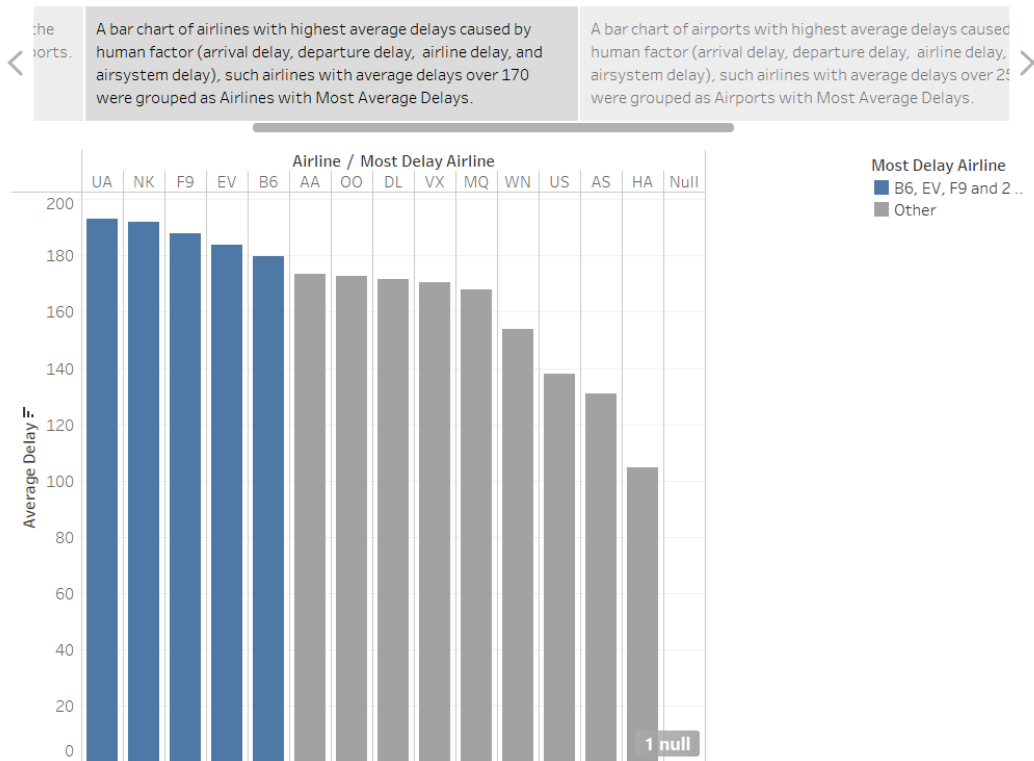
2. Visualization **Two**

For this visualization, a bar plot was picked to compare between the average delay (a quantitative variable) between different airlines, which is a categorical variable. A group of the top 5 airlines with highest average delay was picked, whereas they had an average delay greater than 170.

Two colors were used in the chart: blue, representing the Top 5 Most Delay Airlines, and grey, representing airlines that were not contained in the previously mentioned group.

It was found that UA, NK, F9, EV, and B6 are the Top 5 Most Average Delay Airlines.

US Flights



Three

It was found that ACK, ALO and CIU are the Most Average Delay Airlines.

[illegible]

4. Visualization **Four**

For this visualization, a dashboard was formed combining a line plot and a bar plot.

The multiple line plot chart shows how the aggregated monthly counts of different cancellation reasons escalate (or decline) over time.

The line chart was picked since it's the best suited for seeing how quantitative variables change over time.

It can be seen that the cancellation reason with the highest counts is "Weather", and it has an evident peak in February.

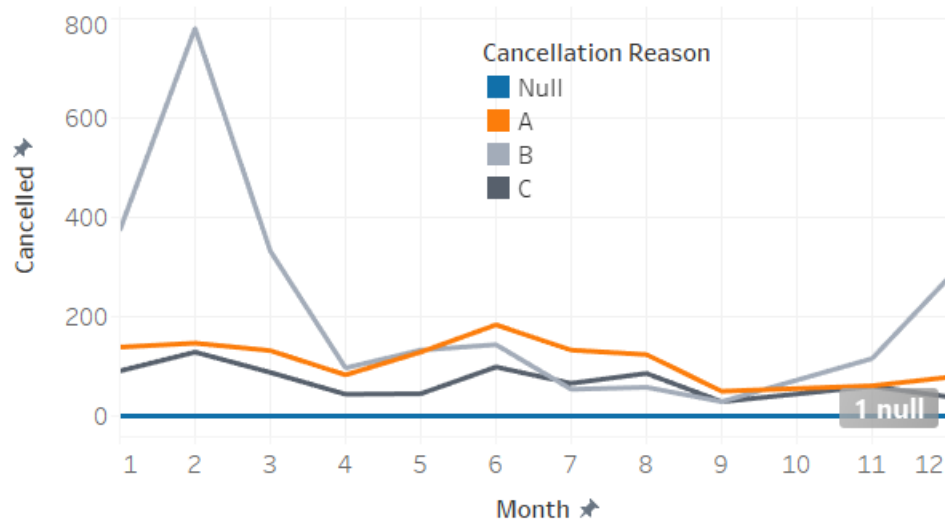
The palette of colors chosen is called "color blind", and as the name suggests, was picked because it uses colors that divert away from red and green to suit viewers that are colorblind.

Colors were used as a mark to differentiate between the line chart of each cancellation reason.

The bar chart here consisted of 2 bars, one denoting the Most Delay Airlines, and the other denotes the remaining airlines.

This bar chart shows how much of the total cancellations were formed by the Most Delay Airlines, and the diagram shows that 35% of the total cancellations occurred due to the Most Delay Airlines.

Cancellations vs. Month of the Year



Most Delayed Airlines Constitute 35% of Total Cancellations

